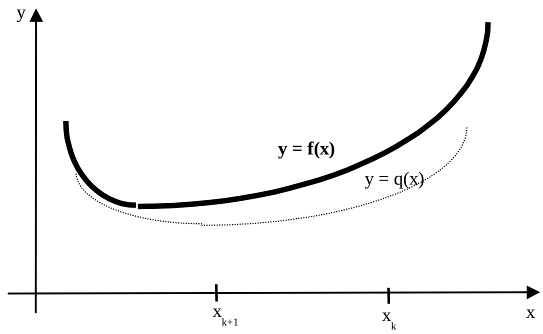
TAUX DE CONVERGENCE D'UNE GENERALISATION DE LA METHODE DE NEWTON : PREMIERE PARTIE

Par Youssef BENADADA

Motivation : La méthode de Newton pour trouver le zéro de l'équation g(x) = 0 peut être facilement généralisée au cas où g est monotone , convexe , mais pas forcément différentiable . La convergence est alors superlinéaire . Le but de cet article est de donner un exemple d'algorithme que l'on peut trouver en recherche opérationnelle et d'étudier sa vitesse de convergence . Dans la deuxième partie de cet article , nous montrerons que la convergence est seulement superlinéaire . [1] ,[2]
Signalons que le lecteur trouvera en fin d'article toutes les définitions nécessaires et références .

1) Introduction : Considérons une fonction f d'une seule variable x à minimiser et supposons qu'au point x_k il est possible d'évaluer $f(x_k)$, $f'(x_k)$ et $f''(x_k)$. Il est donc possible d'approximer f au voisinage de x_k par la fonction :
$q(x) = f(x_k) + (x - x_k)f'(x_k) + \frac{1}{2}(x - x_k)^2 f''(x_k)$. Une estimation de x_{k+1} le minimum de f est donc donnée par le zéro de la dérivée de q : $q'(x_{k+1}) = 0$ ou encore $x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$.

Le processus est illustré sur la figure ci-contre :



La méthode de Newton peut être donc vue comme une technique itérative de résolution des équations de la forme g(x) = 0 .
Considérons le problème suivant :
Trouver $\bar{x} \in I = [a, b]$ tel que $g(\bar{x}) = 0$, où g est une fonction numérique continue strictement monotone sur I telle que : g(a)g(b)<0 . Un tel $\bar{x}$ existe et est unique . Lorsque g est différentiable , la méthode de Newton génère une suite (x_k) de la manière suivante : $x_{k+1} = x_k - \frac{g(x_k)}{g'(x_{k+1})}$ **(1)** .
Si g est deux fois différentiable et $g'(\bar{x}) \neq 0$, alors on peut montrer que la suite (x_k) converge localement vers $\bar{x}$ et que la convergence est quadratique (voir corollaire) . Par ailleurs , si g est convexe ou concave et si $x_k \in I$ pour tout k , alors la convergence est globale .
Supposons maintenant que g est convexe mais non forcément différentiable . Puisque le sous-gradient (voir annexe) est une généralisation de la dérivée pour les fonctions convexes , alors il est naturel de remplacer la relation **(1)** par :
$x_{k+1} = x_k - \frac{g(x_k)}{x_k^*}$ **(2)** où x_k^* est un élément du sous-différentielle $\partial g(x_k)$ de g au point x_k .
Le but de cet article est de faire une étude de la convergence de l'algorithme de Newton généralisé . Nous montrons que la convergence est superlinéaire . Dans un premier temps , il semblait que le taux de convergence était au moins 1.618 . Mais nous verrons ultérieurement dans la seconde partie que ceci est faux : En effet on peut construire des exemples pour lesquels le taux de convergence est α pour tout $\alpha \in]1,2[$.
Dans le reste de cet article , nous supposerons que g est une fonction convexe telle que g(a) < 0< g(b) . Ainsi , g est strictement croissante et continue sur [$\bar{x}$,b] . L'algorithme de Newton peut se résumer comme suit :
Procédure :
(A) On part d'un point initial $x_0 \in [a,b]$ (c'est à dire $g(x_0)>0$) , et on pose k = 0 .
(B) Choisir $x_k^* \in \partial g(x_k)$ et poser $x_{k+1} = x_k - \frac{g(x_k)}{x_k^*}$. On substitue k+1 à k et on retourne à l'étape **(B)** .

2) Convergence superlinéaire : (voir annexe)

Théorème : On considère la suite (x_k) générée par l'algorithme de Newton généralisé pour trouver $\bar{x}$. Alors on a :
a) $\bar{x} \le x_{k+1} \le x_k$, pour tout k entier
b) La convergence de (x_k) vers $\bar{x}$ est superlinéaire .

Démonstration :

a) Par définition de $x_k^* \in \partial g(x_k)$, on a : $g(x_{k+1}) \geq g(x_k) + (x_{k+1} - x_k)x_k^* = 0$ **(3)** , on en déduit que $x_{k+1} \geq \bar{x}$. Par ailleurs , puisque $x_k \geq \bar{x}$, on en déduit que $g(\bar{x}) - g(x_k) - (\bar{x} - x_k)x_k^* \geq 0$, d'où $x_k^* > 0$. Or $g(x_k) \geq 0$, on déduit alors de **(3)** que $x_{k+1} \leq x_k$.

b) Posons maintenant $\bar{t}^* = g_+(\bar{x}) = \mathrm{Max}\{x^* / x^* \in \partial g(\bar{x})\}$. Alors $g(x_k) \geq g(\bar{x}) + (x_k - \bar{x})\bar{x}^* = (x_k - \bar{x})\bar{x}^*$ **(4)** .

Mais d'après **(2)** $g(x_k) = x_k^*(x_k - x_{k+1})$ **(5)** ; en combinant **(4)** et **(5)** , on en déduit que : $0 \leq x_{k+1} - \bar{x} \leq (x_k - \bar{x})(1 - \frac{\bar{x}^*}{x_k^*})$ **(6)** .

Enfin la monotonicité de $\partial g(x)$ adjoint à l'inégalité $\bar{x} \leq x_k \leq x_0$, permet de voir que $0 < \bar{x}^* \leq x_k^* \leq x_0^*$, et la suite (x_k) converge vers $\bar{x}$ au moins linéairement . D'autre part , puisque $\partial g(x)$ est semi-continue supérieurement , et (x_k) converge vers $\bar{x}$ avec $x_k \geq \bar{x}$, alors (x_k^*) converge vers $\bar{x}^*$. La convergence superlinéaire en découle .

Corollaire : on suppose que g est différentiable et que g' est lipchitzienne , alors la convergence de la suite (x_k) générée par la méthode de Newton , est quadratique .

Démonstration :

Sous l'hypothèse de différentiabilité , on peut remplacer dans la relation **(6)** , respectivement x_k^* et $\bar{x}_k^*$ par $g'(x_k)$ et $g'(\bar{x})$. On obtient alors : $0 \leq x_{k+1} - \bar{x} \leq (x_k - \bar{x})\left(1 - \frac{g'(\bar{x})}{g'(x_k)}\right) = (x_k - \bar{x})\frac{g'(x_k) - g'(\bar{x})}{g'(x_k)}$ **(7)** . Or g' est lipchitzienne , donc il existe L > 0 tel que pour tout k entier : $g'(x_k) - g'(\bar{x}) \leq L(x_k - \bar{x})$. En remplaçant dans **(7)** et en utilisant la monotonicité de g , on obtient :

$0 \leq x_{k+1} - \bar{x} \leq (x_k - \bar{x})^2 \frac{L}{g'(\bar{x})}$, d'où la convergence quadratique de la suite (x_k) vers $\bar{x}$.

Remarque : On peut obtenir le même taux de convergence sans pour autant que g soit différentiable . En effet , supposons qu'il existe L > 0 tel que pour tout k entier : $0 \leq x_k^* - \bar{x}^* \leq L(x_k - \bar{x})$; alors la relation **(6)** devient : $0 \leq x_{k+1} - \bar{x} \leq (x_k - \bar{x})^2 \frac{L}{\bar{x}^*}$.

On en déduit donc la convergence quadratique de la suite (x_k) vers $\bar{x}$.

ANNEXE

Définition 1 : Soient Γ une multiapplication d'un espace topologique X dans un espace topologique Y , et $x_0 \in X$; on dit que Γ est **semi-continue supérieurement** en x_0 si pour tout ouvert ω contenant $\Gamma(x_0)$, il existe un voisinage Ω de x_0 tel que : $x \in \Omega \Rightarrow \Gamma(x) \in \omega$. On dit que Γ est **semi-continue supérieurement** si elle l'est en tout point de X , et si , de plus $\Gamma(x)$ est un ensemble compact pour tout x appartenant à X .

Définition 2 : Soit $\theta : \mathbb{R} \to \mathbb{R}$ une fonction convexe ; alors x* est un **sous-gradient** de θ au point x si : $\theta(y) \geq \theta(x) + (y-x)x^*$, pour tout $y \in \mathbb{R}$. On désigne alors par $\partial\theta(x)$ l'ensemble des sous-gradients de θ au point x . $\partial\theta(x)$ est appelé le **sous-différentiel** de θ au point x . On démontre alors que la multiapplication $x \to \partial\theta(x)$ est semi-continue supérieurement au sens au sens des multiapplications . Remarquons également que $\partial\theta(x) = [\theta_-'(x), \theta_+'(x)]$ avec :

$$\theta_-'(x) = \lim_{\substack{y \to x \\ y < x}} \frac{\theta(y) - \theta(x)}{y - x} = \min[x^* / x^* \in \partial\theta(x)] \text{ et } \theta_+'(x) = \lim_{\substack{y \to x \\ y > x}} \frac{\theta(y) - \theta(x)}{y - x} = \max[x^* / x^* \in \partial\theta(x)] .$$

Définition 3 : [4,5]

Considérons la suite (x_k) qui converge vers $\bar{x}$. L'ordre de convergence de (x_k) est défini comme étant la borne supérieure des réels p strictement positif satisfaisant : $0 \leq \limsup_{k \to \infty} \frac{|x_{k+1} - \bar{x}|}{|x_k - \bar{x}|^p} < \infty$.

Exemple 1 : La suite $x_k = a^k$, avec 0<a<1 converge vers 0 avec un ordre de convergence égal à 1 puisque $\frac{x_{k+1}}{x_k} = a$.

Exemple 2 : La suite $x_k = a^{2^k}$, avec 0<a<1 converge vers 0 avec un ordre de convergence égal à 2 puisque $\frac{x_{k+1}}{x_k^2} = 1$.

La classe des algorithmes ayant un ordre de convergence égal à 1 est très importante ; nous allons en traiter des cas particuliers :

Définition 4 : Si la suite (x_k) converge vers $\overline{x}$ avec $\lim_{k\to\infty} \frac{|x_{k+1} - \overline{x}|}{|x_k - \overline{x}|} = \beta < 1$, alors on dit que cette suite converge **linéairement** vers $\overline{x}$ avec le taux de convergence égal à β . si en plus $\beta = 0$, alors on dit que (x_k) converge **superlinéairement** vers $\overline{x}$.

A suivre

Bibliographie :

[1] Y. BENADADA (1989) Approches de résolution du problème de programmation fractionnaire généralisée . Thèse de Doctorat , Université de Montréal (Montréal , CANADA) .

[2] Y. BENADADA , J-P CROUZEIX , J-A FERLAND (1993) Rate of convergence of a generalization of Newton's Méthod . Journal of optimization theory and application .Vol 78 , n°3 , Septembre 1993 .

[3] C. BERGE (1966)
Espaces topologiques : fonctions multivoques . Deuxième édition . Dunod .

[4] M. MINOUX
Programmation Mathématique , théorie et algorithme . Tome 1 .

[5] D.G LUENBERGER (1984)
Linear and nonlinear programming . Deuxième édition , Addison Wesley .